E83-10092

CR-169602

NOV 8 - 1982

AgRISTARS

Yield Model Development

YM-J2-04359
JSC-18573

A Joint Program for Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing

September 1982

A COMPARISON OF MEASURED AND ESTIMATED METEOROLOGICAL DATA FOR USE IN CROP GROWTH MODELING

(E83-10092) A COMPARISON OF MEASURED AND ESTIMATED METEOROLOGICAL DATA FOR USE IN CROP GROWTH MODELING (Department of Agriculture) 5 p HC A02/MF A01 CSCL 02C

N83-16808

Unclas
G3/43 00092

PREPRINT FOR THE ANNUAL MEETING OF THE AMERICAN STATISTICAL ASSOCIATION
– BUSINESS AND ECONOMIC STATISTICS SECTION –
CINCINNATI, OHIO, AUGUST 16–19, 1982

Lyndon B. Johnson Space Center
Houston, Texas 77058

ORIGINAL PAGE IS
OF POOR QUALITY

A COMPARISON OF MEASURED AND ESTIMATED METEOROLOGICAL DATA FOR USE IN CROP GROWTH MODELING[1]

Charles R. Perry, Jr., U.S. Department of Agriculture, Statistical Reporting Service[2]

Julian L. Rogers, U.S. Department of Agriculture, Federal Crop Insurance Corporation[2]

Joe T. Ritchie, U.S. Department of Agriculture, Agriculture Research Service[3]

ABSTRACT

Modern crop growth simulation and crop condition assessment models require daily input of maximum temperature, minimum temperature, precipitation, and solar radiation data. Gridded spatial estimates of these variables are prepared for agricultural use from World Meteorological Organization surface reports and enhanced by polar orbiting satellites. If sufficiently accurate, these data which are made available from an operational system soon after they are prepared, may reduce the cost of reliably estimating world crop conditions. The gridded estimates were compared with daily meteorological data measured at various agricultural research facilities across the United States to determine their level of accuracy. Preliminary results indicate that daily maximum temperature can be determined to within 9.1 degrees Celsius with ninety percent confidence. With similar levels of confidence, daily minimum temperature can be determined to within 6.7 degrees Celsius, daily solar radiation to within 231.2 cal/cm^{2}min, and daily precipitation to within 9.7 millimeters.

1. INTRODUCTION

This paper outlines a study conducted by the Yield Research Branch of the Statistical Reporting Service as part of the Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing (AgRISTARS) program, a joint USDA, NASA, NOAA, USDC, and AID research effort to determine the feasibility of integrating aerospace remote sensing technology into existing or future USDA data acquisition systems. The purpose of this study was to compare the U.S. Air Force (USAF) Agromet data to measured daily meteorological data collected at various agricultural research facilities across the United States. The daily data elements evaluated were maximum temperature, minimum temperature, total precipitation, and solar radiation. Measured data are not readily available for evaluation of potential evapotranspiration and crop growth simulation. If the Agromet data--which are routinely available and prepared for agricultural use from World Meteorological Organization (WMO) surface reports and enhanced by polar orbiting satellites--are accurate enough for plant and soil water modeling, data collection costs may be significantly reduced.

2. BACKGROUND

Interest in worldwide crop production and economic conditions has grown in recent years. In response, the USAF has developed a complex model at Global Weather Central (GWC), Offutt Air Force Base, Nebraska, to provide daily meteorological data specifically tailored for agricultural use. Data to provide Agromet information comes from the WMO network surface reports and polar orbiting satellites measuring reflectance, radiance, and temperature. An automated cloud analysis model (3DNEPH) estimates the effect of clouds on the radiation balance (refs. 1 and 2).

Through an agreement between the USAF and the U. S. Department of Agriculture (USDA), these Agromet data are available for crop condition assessment and research use on a real time basis. No assessment of the accuracy of these data over a season has been done.

3. DESCRIPTION OF DATA

The data elements described include only those for which an evaluation is being done. Agromet uses the 1/8 mesh AF GWC grid on a polar stereographic projection, which gives approximately 25-nm grid point spacing at 60° N, and all data elements are provided on this grid point basis. These estimates are currently provided for most of the United States and many areas of the Northern Hemisphere. Preparation of estimates for the United States began in June 1981.

3.1 MAXIMUM AND MINIMUM TEMPERATURES

Maximum and minimum temperatures are estimated every 3 hours from satellite temperature estimates. The highest and lowest estimates are saved for use at the end of the day. Surface reports (three hourly temperatures and daily reported maximum-minimums) are used in an analysis with the satellite temperature estimates to create maximum-minimum temperatures for each grid point in each geographic region.

3.2 PRECIPITATION

Precipitation reports from surface stations are each assigned to the nearest grid point. Reported values are accumulated along with estimated amounts based on weather conditions from surface reports. Daily accumulations are

made of the greater of the above amounts. Checks are made for convective precipitation, and one-half of the daily accumulation is spread to adjacent grid points if no other precipitation is reported for those grids. Spreading is based on estimated amounts from the 3DNEPH cloud analysis. A ratio of reported to estimated precipitation (R/E) is calculated for each grid point and spread using a linear distance weighting.

The reported precipitation is used for those points for which it is available. At other points the R/E is used to determine a value. In a few cases the quantitative precipitation forecast (QPF) from the 3DNEPH is used. This is generally less than 12 grid points for an area.

3.3 SOLAR RADIATION

Clear sky direct solar radiation is calculated from long standing, well-known equations. The clear sky solar radiation is adjusted for cloudiness using 15 cloud layers in the 3DNEPH cloud model. A detailed explanation of Agromet net solar radiation computation is given in USDA ETAC/TN-81-001, March 1981.

3.4 OTHER AGROMET DATA

Other data elements produced by Agromet are not being evaluated at this time. A description of these data can be obtained from ETAC/TN-81-001.

4. DATA COLLECTION

4.1 AIR FORCE AGROMET DATA

The Agromet data being used in this study were processed in the Foreign Agricultural Service (FAS), Foreign Crop Condition Assessment Division (FCCAD), Houston, Texas.

Agromet data are processed continuously (every 3 hours) at GWC, Offutt Air Force Base, Nebraska. Each day at 2400 GMT the daily Agromet Data Summary is prepared. Data from these daily reports are assembled every Monday into a weekly (Monday through Sunday) data set, and a magnetic tape is Air Expressed to FCCAD, Houston, Texas. The data are processed each Tuesday and are available for operational use generally within 60 hours after preparation at GWC. Data for comparison to measured data have been extracted from the FCCAD disk file for use in statistical analysis.

Grid elements for each grid cell in which a research location is situated were compared against the corresponding measured data element for the same day.

4.2 MEASURED DATA

The ground-measured meteorological data were collected from routine measurements made at various agricultural research centers in the United States and were assembled by the USDA/Agricultural Research Service (ARS), Crop Systems Evaluation Unit at Temple, Texas.

Daily measurements included maximum-minimum temperatures, precipitation, and solar radiation.

Both of these data sets were furnished to the Yield Model Development (YMD) project, Houston, for use in statistical evaluation of the Agromet data.

5. STATISTICAL ANALYSIS

Analytic and graphical methods were used to characterize the error structure of the Agromet estimators. For each variable the mean, variance, and other standard statistics were computed for differences between the Agromet measurements and the station measurements. The mean and variance of the differences provide estimates of the bias and variance expected in using the Agromet data, assuming that the difference estimates the error made by using the Agromet data instead of the station data. Tests for month and station effects were examined using a two-factor linear model with month and station random factors. Similar nonparametric tests were performed and the results compared with the parametric results. Histograms and time plots were made and visually examined for each set of differences. These graphs provided insight into the expected behavior of the estimators.

Since the measured data provides values at a specific point and the Agromet data provides estimates over a large area (no smaller than 25 by 25 nm), differences in the two data values were expected. Understanding these differences is necessary before using the Agromet data for large-area agricultural estimation without regard to ground station information.

6. SOME PRELIMINARY RESULTS

One way to characterize the errors made in using the daily Agromet cell estimates as an estimator of the station point measurements is to fit a general linear model to the daily differences between these measurements. Since the station locations are a subset (more or less randomly located in the major agricultural land of the United States) of a much larger set of points for which we wish to make inferences the classification variable station location was considered a random effect. How to treat the classification variable month was not so obvious. The months themselves were not a random sample of all months; they were the twelve consecutive months for which the data were gathered. However, the effect on the differences was considered a random effect because we were not interested in comparing specific months; we were interested only in ascertaining how much of the estimation error was attributable to month-to-month changes in the differences. The analysis for the variable daily maximum temperature is presented as an example of the technique used and the results obtained.

The two-factor random-effects model was fit to the data. The components of variance associated with the residual error, the month effect, the station effect, and the station

ORIGINAL PAGE IS
OF POOR QUALITY

and month interaction were estimated by three standard methods (ref 3, pp. 433), and hypotheses were tested using several analogous type sums of squares to ascertain if any of the components of variance could reasonably be considered to be zero (refs 4 and 5). The results for testing the hypothesis that a variance component is zero are summarized in Table 1; the results are given using type IV analogous sums of squares. The results were identical when other type sums of square were used. Table 2 gives the estimates for the variance components for each of the three methods employed. The conclusion is that the variance component associated with the classification variable month should be considered zero. Following standard statistical procedure, the classification variable month was dropped from the analysis and the reduced model refitted to the data, the hypotheses retested, and the variance components reestimated. The results are presented in Tables 3 and 4.

The preliminary conclusion from the results given in Tables 1-4 is that the variance components associated with the station location and the station month interaction are not zero and that both contribute significantly to the total error made in using the Agromet daily cell maximum-temperature estimate as an estimator of the associated station point maximum-temperature measurement. There is no reason to suspect that the variance component associated with the classification variable month contributes significantly to the overall error structure; hence, it is reasonable to consider it zero.

From Table 4, the residual variance, the station variance, and the station*month variance contribute respectively about 85 percent, 9 percent, and 7 percent to the total variance. This implies that about 90 percent of the time the daily cell grid maximum temperature estimate is within 9.1 degrees Celsius (9.1=1.64 times the estimated standard deviation) of the associated point station measurement for daily maximum temperature; the contribution of the residual error to this value is 8.4 degrees Celsuis.

Similar results indicate that daily minimum temperature can be determined to within 6.7 degrees Celsius with 90 percent confidence. With similar levels of confidence, daily solar radiation can be determined to within 231.2 (calories/cm^{2}min) and daily precipitation to within 9.7 millimeters. For daily minimum temperature, the residual variance, the station variance, and the station*month variance were estimated to contribute respectively 75 percent, 12 percent, and 13 percent to the total variance. For daily solar radiation, the residual variance, the station variance, and the station*month variance were estimated to contribute respectively 96 percent, 2 percent, and 2 percent to the total variance. For daily precipitation, the residual variance and the station variance were estimated to contribute respectively 99 percent, and 1 percent to the total variance; the variance component associated with the station*month interaction was not significant. A complete discussion of these analyses can be found in Perry 1982 (ref. 6).

[1] Contribution of the Yield Model Development (YMD) project within the Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing (AgRISTARS) program, a joint program of USDA, USDC, NASA, USDI, and AID.

[2] AgRISTARS, 1050 Bay Area Blvd., Houston, Texas 77058.

[3] Grasslands, Soil and Water Laboratory, Temple, Texas 76503

REFERENCES

1. USAF ETAC/TN81/001 Soil Moisture and Agromet Models.

2. Gridded Meteorological Data Extraction System, Users Guide.

3. SAS Users Guide 1979 Edition, SAS Institute, Cary, North Carolina

4. R. T. Freund and R. C. Littel. SAS for Linear Models. SAS Institute, Cary, North Carolina

5. S. R. Searle: Linear Models. Wiley, (New York), 1971.

6. C. R. Perry, J. L. Rogers, and J. T. Ritchie: An Analysis of the Differences Between Gridded Agromet Meteorological Cell Estimates and Station Meteorological Point Measurements, AgRISTARS Document YM-U2-04355, JSC-18569. To be published.

TABLE 1.- ANALYSIS OF VARIANCE FOR THE FULL MODEL USING TYPE IV SUMS OF SQUARES TO TEST

[H_0: VAR(MONTH) = 0; H_0: VAR(STATION) = 0; H_0: VAR(STATION*MONTH) = 0]

SOURCE OF VARIATION	DEGREES OF FREEDOM	SUM OF SQUARES	COMPUTED F_0	PROB F > F_0	EXPECTED MEAN SQUARE
MONTH	11	704.5	0.67	0.76	$\sigma_E^2 + 29.15\sigma_{S*M}^2 + 297.97\,\sigma_M^2$
STATION	10	9712.5	10.18	0.0001	$\sigma_E^2 + 28.89\,\sigma_{S*M}^2 + 322.41\,\sigma_S^2$
STATION*MONTH	104	9921.4	3.67	0.0001	$\sigma_E^2 + 29.15\,\sigma_{S*M}^2$
ERROR	3551	92400.9			σ_E^2

TABLE 2.- VARIANCE COMPONENT ESTIMATES FOR THE FULL MODEL

VARIANCE COMPONENT	VARIANCE COMPONENT ESTIMATION PROCEDURE		
	TYPE I SS	MIVQUEO	MAXIMUM LIKELIHOOD
VAR (MONTH)	-0.12	-0.12	0.00
VAR (STATION)	2.68	2.69	2.43
VAR (STATION*MONTH)	2.38	2.39	2.23
VAR (ERROR)	26.02	26.00	26.02

TABLE 3.- ANALYSIS OF VARIANCE FOR THE REDUCED MODEL USING TYPE IV SUMS OF SQUARES TO TEST
[H_0: VAR(STATION) = 0 ; H_0: VAR(STATION*MONTH) = 0]

SOURCE OF VARIATION	DEGREES OF FREEDOM	SUM OF SQUARES	COMPUTED F_0	PROB F > F_0	EXPECTED MEAN SQUARES
STATION	10	9897.71	10.83	0.0001	$\sigma_E^2 + 28.89\,\sigma_{S*M}^2 + 330.39\,\sigma_S^2$
STATION*MONTH	104	10509.79	3.51	0.0001	$\sigma_E^2 + 29.16\,\sigma_{S*M}^2$
ERROR	3551	92400.94			σ_E^2

TABLE 4.- VARIANCE COMPONENT ESTIMATES FOR THE REDUCED MODEL

VARIANCE COMPONENT	VARIANCE COMPONENT ESTIMATION PROCEDURE					
	TYPE I SS		MIVQUEO		MAXIMUM LIKELIHOOD	
	ESTIMATE	PERCENTAGE	ESTIMATE	PERCENTAGE	ESTIMATE	PERCENTAGE
VAR (STATION)	2.71	8.8	2.70	8.7	2.43	7.9
VAR (STATION*MONTH)	2.24	7.2	2.27	7.3	2.23	7.3
VAR (ERROR)	26.02	84.0	26.00	84.0	26.02	84.8
TOTAL VARIANCE	30.97	100.0	30.97	100.0	30.68	100.0